

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

VOICE ACTIVATED, VOICE RESPONSIVE PRODUCT LOCATOR SYSTEM,  
INCLUDING PRODUCT LOCATOR METHOD UTILIZING PRODUCT BAR CODE AND  
PRODUCT-SITUATED, LOCATION-IDENTIFYING BAR CODE

Inventor: Kenneth P. Glynn

Jerome R. Mahoney

Assignee: iVoice.com, Inc.

Attorney Docket No. IVC-107C

This patent application is a continuation-in-part of United States copending patent application Serial Number 09/653,658 filed on August 31, 2000 entitled "Voice Activated/ Voice Responsive Item Locator", assigned to the same assignee as designated herein and having Jerome R. Mahoney as a common inventor.

Kenneth P. Glynn, Esq.  
Attorney for Applicant  
Reg. No. 26,893  
Glynn and Associates  
24 Mine St.  
Flemington, NJ 08822-1731  
tele (908) 788-0077  
fax (908) 788-3999

0070112100KPG-hd

5 VOICE ACTIVATED, VOICE RESPONSIVE PRODUCT LOCATOR  
SYSTEM, INCLUDING PRODUCT LOCATION METHOD  
UTILIZING PRODUCT BAR CODE AND PRODUCT-SITUATED,  
LOCATION-IDENTIFYING BAR CODE

(Attorney Docket No: IVC-107C)

10 REFERENCES TO RELATED APPLICATIONS

15 This patent application is a continuation  
-in-part of United States copending patent  
20 application Serial Number 09/653,658 filed on  
August 31, 2000 and entitled "Voice  
25 Activated/Voice Responsive Item Locator",  
assigned to the same assignee as designated  
30 herein and having Jerome R. Mahoney as a common  
35 inventor.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to voice

activated/voice responsive item locators, i.e.  
item directories, which direct a user such as a  
consumer or shopper, to a specific location to  
view, retrieve, order, purchase or otherwise use  
5 the information obtained in the system. Further,  
the present invention includes within the  
aforesaid system, a method of collecting location  
data for the system which involves the use of  
product-situated, product-identifying bar codes  
10 and product-situated, location-identifying bar  
codes. These are read sequentially and stored in  
the main processor of the system to provide  
location information to subsequent users.  
Typically, the present invention could be used at  
15 retail stores to locate items to be purchased.  
Alternatively, it could be used at a production

facility or distribution facility having a large number of parts, to locate specific parts for as needed. In other embodiments, it could be used in non-commercial entities, such as public libraries to locate a particular book. The locator of the present invention relies upon a specific software module to accomplish voice recognition and response, and includes manager programming for customization, updates and modifications.

10      2. Information Disclosure Statement

15      The state of the art for acquiring product location information involves the use of manually collected, inputted data. Bar codes have been used for years to identify products, but not to identify locations.

The following prior art patents represent

stored in a memory, in response to the  
information of recognition result from the speech  
recognition unit, and a circuit that prevents  
transmission of signals from the telephone  
5 network to the receiver when the regenerated  
speech information is sent to the receiver.  
Furthermore, it is desirable for this device to  
be provided with a circuit that prevents  
generation of ringing tones when an incoming call  
10 arrives.

U.S. Patent No. 5,136,634 to David C. Rae et  
al. describes voice operated facsimile machine  
network which includes a method and apparatus for  
transmitting specifically requested graphic  
15 and/or textual data from an unattended database  
storage location to a requestor's facsimile.

machine over a telephone line which includes a  
host computer such as a PC modified with a  
facsimile transmission board and a voice  
generation board. The host computer receives  
5 incoming phone calls and prompts the caller using  
the voice board to select data files by using the  
DTMF keys of a standard telephone handset. The  
PC can be left unattended and can run  
automatically in the facsimile transmission mode.  
10 Callers can immediately access needed textual and  
image data with the use of just a standard  
telephone and facsimile machine. Multiple  
workstation nodes can be configured in a network  
setup to handle a high volume of calls in real  
15 time and to allow multiple data services to  
operate simultaneously.

U.S. Patent No. 5,165,095 to Mark A.

Borcherding describes a method for dialing a  
telephone, using voice recognition to initiate  
the dialing and to determine the correct  
5 telephone number. The dialing is initiated with  
a spoken dial command that is recognized by using  
speaker independent templates that are stored  
locally with respect to the caller's telephone.  
The correct telephone number is recognized by  
10 using speaker dependent template that are  
downloaded from a central database or by using  
speaker independent templates stored locally.

U.S. Patent No. 5,168,548 to Steven Kaufman  
et al. describes a reporting system which is  
15 disclosed herein, a speech recognizer which is  
used to select selections of text from a report

form stored in a computer and to insert  
recognized terms in the text thereby to generate  
a report text under voice control. A command  
interpreter, also responsive to spoken words,  
5 initiates creation of the report text and its  
subsequent storing, printing and transmission.  
The command processor is responsive to respective  
spoken commands to select a destination telephone  
number and to cause the report text to be sent to  
10 apparatus for converting report text to image  
data and for modulating an audio band signal with  
the image data for facsimile transmission over  
telephone lines.

U.S. Patent No. 5,222,121 to Keiko Shimada  
15 describes a voice recognition dialing unit of a  
telephone mounted on a vehicle or similar mobile



body and which allows a call to be originated  
with ease. When the user of the telephone enters  
a voice command on voice inputting section, the  
dialing unit originates a call automatically and  
5 thereby connects the other party to the telephone  
line. In a call origination procedure, the  
operations for call origination and the  
verifications are performed between the user and  
the unit in an interactive sequence. In a  
10 preferred embodiment, the unit has a particular  
call origination procedure in which, when the  
other party recognized by the unit is wrong as  
determined by the user by verification, lower  
place candidates for the other party are called  
15 up in response to a particular voice command. In  
an alternative embodiment, the unit indicates the

other party by voicing a name for verification  
purpose. The alternative embodiment selects and  
stores only the name of the other party in  
response to an entered voice signal and, in the  
5 event of response for verification, combines the  
name having been stored and response information  
stored beforehand to produce composite response  
voice.

U.S. Patent No. 5,231,670 to Richard S.

10 Goldhor et al. describes a system and method for  
generating text from a voice input that divides  
the processing of each speech event into a  
dictation event and a text event. Each dictation  
event handles the processing of data relating to  
15 the input into the system, and each text event  
deals with the generation of text from the

inputted voice signals. In order to easily distinguish the dictation events from each other and text events from each other the system and method creates a data structure for storing certain information relating to each individual event. Such data structures enable the system and method to process both simple spoken words as well as spoken commands and to provide the necessary text generation in response to the spoken words or to execute an appropriate function in response to a command. Speech recognition includes the ability to distinguish between dictation text and commands.

U.S. Patent No. 5,239,586 to Kuniyoshi Marui describes a voice recognition system which comprises a handset and a hands-free microphone

for generating an input audio signal, a high-pass  
filter for eliminating low frequency components  
from the signal from the handset or hands-free  
microphone, a signal level controller for  
5 adjusting the level of the high-pass signal in  
response to the user of either the handset or  
hands-free microphone, a storer for storing the  
speech data and a controller for controlling the  
storer so that a user's utterance is stored or  
10 the user's utterance is recognized by comparing  
the utterance to speech data already stored. The  
handset hook switch provides an on-hook control  
signal to reduce amplifier gain during hands-free  
microphone operation.

15 U.S. Patent No. 5,301,227 to Shoichi Kamei  
et al. describes an automatic dial telephone that

is useable in a motor vehicle, when a voice input  
is provided during a period in which input of the  
names of called parties is awaited, a voice  
pattern of the name of the called party is  
5 compared with reference patterns of called  
parties stored in reference patterns storing  
device, to determine the degree of the similarity  
therebetween. The names of the called parties  
are output to a user in the order of decreasing  
10 degree of similarity. Each time the name of a  
called party is output, a command word for  
confirmation is awaited from a user for a  
predetermined time period. When a voice  
confirmation command is input and is recognized  
15 during this waiting period, a telephone number  
corresponding to the name of the called party is

supplied to a channel. Consequently, the command word for confirmation may be input only if the name of the called party outputted is one desired by the user. Sensors continually monitor the driving condition of the motor vehicle in which the telephone is installed. When the operation of the steering wheel or brakes of the motor vehicle exceeds a predetermined threshold or the speed of the motor vehicle is excessive, the sensors generate safety signals that inhibit the operation of the telephone.

U.S. Patent No. 5,335,276 to E. Earle Thompson et al. describes a communication system which is provided with multiple purpose personal communication devices. Each communication device includes a touch-sensitive visual display to

communicate text and graphic information to and  
from the user and for operating the communication  
device. Voice activation and voice control  
capabilities are included within communication  
5 devices to perform the same functions as the  
touch-sensitive visual display. The  
communication device includes a built-in modem,  
audio input and output, telephone jacks and  
wireless communication. A plurality of  
10 application modules are used with personal  
communication devices to perform a wide variety  
of communication functions such as information  
retrievable, on-line data base services,  
electronic and voice mail. Communication devices  
15 and application modules cooperate to allow  
integrating multiple functions such as real time

communication, information storage and  
processing, specialized information services, and  
remote control of other equipment into an  
intuitively user friendly apparatus. The system  
5 includes both desktop and hand-held communication  
devices with the same full range of communication  
capabilities provided in each type of  
communication device.

U.S. Patent No. 5,349,636 to Roberto

10 Irribarren describes a communication system for  
verbal telephonic communication which has a voice  
message system for storing and retrieving voice  
messages integrated with a computer database  
accessing system for storing and retrieving text  
15 messages from a separate computer system and for  
converting the text messages into voice. The



systems are integrated via a network which coordinates the functions of each individual system. Additionally, the input/output ports of the voice message system and the computer

5 database accessing system are connected in a parallel fashion to at least one telephone line.

In this configuration a user may access both voice messages and database information, including text or electronic mail messages, with  
10 a single telephone call. Optionally, facsimile messages can be stored, retrieved and manipulated with a single telephone call.

U.S. Patent No. 5,406,618 to Stephen B. Knuth et al. describes a telephone answering  
15 device that is activated by a proximity sensor when a user crosses its field of detection and

whose operation is controlled by simple voice commands. The device incorporates speaker-independent voice recognition circuitry to respond to spoken commands of the user that are elicited by a system generated voice request menu. The telephone answering device performs all the basic functions of a telephone answering machine in response to these simple commands and there is no need for the user to manually operate the telephone answering device.

U.S. Patent No. 5,602,963 to W. Michael Bissonnette et al. describes a small, portable, hand-held electronic personal organizer which performs voice recognition on words spoken by a user to input data into the organizer and records voice messages from the user. The spoken words

and the voice messages are input via a  
microphone. The voice messages are compressed  
before being converted into digital signals for  
storage. The stored digital voice messages are  
5 reconverted into analog signals and then expanded  
for reproduction using a speaker. The organizer  
is capable of a number of different functions,  
including voice training, memo record, reminder,  
manual reminder, timer setting, message review,  
10 waiting message, calendar, phone group select,  
number retrieval, add phone number, security and  
"no" logic. During such various functions, data  
is principally entered by voice and occasionally  
through use of a limited keypad, and voice  
15 recordings are made and played back as  
appropriate. A visual display provides feedback

to the user. During the various function, the user can edit various different data within the organizer by eliminating or correcting such data or entering new data.

5

U.S. Patent No. 5,621,658 to Brion K.

Jackson describes an action contained within an electronic mail object which is communicated from a data processing system to another data processing system via an audio device. The

10

action is executable on a data processing system.

At the sending data processing system, the action is converted to a predetermined audio pattern.

The electronic mail object may contain text in addition to an action. The text is also

15

converted to an audio pattern. The audio patterns are then communicated to the audio

device over telephone lines or other  
communication medium. At the receiving end, the  
audio device records the object. A user can  
provide the recorded object to a data processing  
5 system, which then executes the action and  
converts the text audio patterns back to text.  
In addition, the action can be converted to text  
and displayed on the data processing system.

U.S. Patent No. 5,631,745 to John J. Wong et  
10 al. describes a telephone terminal adapted for  
business or home use that includes the ability to  
receive and send facsimiles, a voice answering  
function and a computer modem. Various input and  
output devices may be used for the facsimile  
15 function. A voice annotated facsimile may be  
sent and received. At the same time the

facsimile is viewed on a video monitor or  
ordinary television set, an accompanying voice  
message is heard through the sound system of the  
monitor or television set. The terminal has an  
5 architecture including a central processor and an  
internal bus structure to which several types of  
memory, various input-output devices and an  
interface with the telephone line are connected,  
among others. Audio Random Access Memory (ARAM)  
10 is used for storing both facsimile data and voice  
data.

U.S. Patent No. 5,671,328 to Gregory P.  
Fitzpatrick et al. describes a method and data  
processing system which are disclosed for  
15 automatically creating voice processing template  
entries. In one embodiment, the invention

automatically assembles a plurality of commands  
received by the data processing system, at least  
one of said commands having a voice recognition  
criteria component associated therewith, counts  
5 the occurrences of the plurality of commands,  
assembles voice recognition criteria components  
associated with the plurality of commands, and,  
as a result of the occurrence count exceeding a  
predefined minimum, constructs a voice  
10 recognition template entry by associating the  
assembled voice recognition criteria components  
with the assembled plurality of commands.

U.S. Patent No. 5,850,627 to Joel M. Gould  
et al. describes a word recognition system which  
15 can: respond to the input of a character string  
from a user by limiting the words it will

recognize to words having a related, but not necessarily the same, string; score signals generated after a user has been prompted to generate a given word against words other than the prompted word to determine if the signal should be used to train the prompted word; vary the number of signals a user is prompted to generate to train a given word as a function of how well the training signals score against each other or prior models for the prompted word; create a new acoustic model of a phrase by concatenating prior acoustic models of the words in the phrase; obtain information from another program running on the same computer, such as its commands or the context of text being entered into it, and use that information to vary which



words it can recognize; determine which program  
unit, such as an application program or dialog  
box, currently has input focus on its computer  
and create a vocabulary state associated with  
5 that program unit into which vocabulary words  
which will be made active when that program group  
has the focus can be put; detect the available  
computational resources and alter the  
instructions it executes in response; test if its  
10 ability to respond to voice input has been shut  
off without user confirmation, and, if so, turn  
that ability back on and prompt the user to  
confirm if that ability is to be turned off;  
store both a first and a second set of models for  
15 individual vocabulary words and enable a user to  
selectively cause the recognizer to disregard the

second set of models for a selected word; and/or  
score a signal representing a given word against  
models for that word from different word model  
sets to select which model should be used for  
5 future recognition.

Notwithstanding the prior art, the present  
invention is neither taught nor rendered obvious  
thereby.

#### SUMMARY OF THE INVENTION

10 A voice activated/voice responsive item  
locator system is disclosed to enable a user to  
speak into the system and have the system respond  
with location information for an item requested  
by the user. For example, shopper at a home  
15 supply store may pick up a locator phone or just  
speak into a wall mounted or otherwise situated

microphone and say "Locate Outdoor Paint" or  
"Find Hammers" or simply state what is sought  
without the use of a verb, e.g. "Caulking". The  
system may reply either with voice or visual  
5 (words on a screen, or map), or both voice and  
visual, e.g. "Aisle 3, Shelf 4". In some  
instances the system will reply, for example,  
with a "Repeat", or "Restate in different words"  
or "Please talk to information desk" or other  
10 default instructions.

The present invention also includes a  
method of creating data for locating items so  
that the system is efficiently loaded with  
location data both prior to use by the customers  
15 or other users, as well as so that the system may  
be updated as desired while it is in use. This

method involves utilization of bar codes to  
determine item identity, and the use of separate  
bar codes to determine locations. These separate  
location bar codes are physically located on the  
5 products themselves, e. g., at least one sample  
of a set of items will contain the location bar  
code in addition to its own product-identifying  
bar code. This location data is read in  
conjunction with item identification data by bar  
10 code readers, fed to a processor in a  
recognizable combined format, and then stored and  
used as the resource data of the locator system.

For example, a supermarket could assign  
unique bar codes to each aisle, create bar code  
15 labels and attach them one or more units of each  
particular type of item using the appropriate

aisle code, and then program the system according to the following simple process:

a) The processor will be programmed to read and identify products by the universal price code ("UPC") inputs from a bar code reader, and will likewise be programmed to recognize and identify locations by bar code inputs from a bar code reader, that is, the processor will be programmed to understand the codes created for particular locations to be included in the supermarket product location system;

b) The processor will also be programmed to couple items (products) to locations when read in sequence. In other words, when a reader inputs a UPC for an item and then reads the location bar code on it, this tells the processor to create a

matching set of pairs of products and locations  
for all products read. In an alternative  
embodiment, each type of item could be read after  
the location reading to create location data  
5 pairings. In other embodiments, the bar code  
reader need not be a portable reader operated for  
the sole purpose of gathering the location data.  
For example, the reading could take place during  
inventory checking or at the cash register itself  
10 during checkou by the customers. In one preferred  
embodiment, the readings and even the attachments  
of the location bar codes to the products, is  
performed at t eh time of stocking the items and  
then the system is kept current and modified  
15 through checkout readings.

The overall locator system may be a stand

alone device, but in most embodiments would be part of an internal connected system. It could be an intranet or secured internet system, but would in many cases be a storewide system with a plurality of user locations (units, phones, or microphones, with feedback at each location). The system will include an embedded voice-driven interface for speech control of: (1) operational instructions; (2) core system locator function operations, that is, recognition of specific requests and responses thereto; and, (3) optional and default functions. In preferred embodiments, the present invention device is both operated by speech (speech or voice activated) and speech responsive (voice answers and instructions to the user from the system). Thus, the present

invention device relies upon automatic speech  
recognition (ASR), either in place of or in  
addition to manual locator systems, e.g. book,  
list, map and computer directories. In some  
5       embodiments, user feedback features are included  
wherein both audio and visual feedback is given  
to a user in response to recognizable voice  
signals, while in other possible embodiments, the  
user may designate audio or visual.

10

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention should be more fully  
understood when the specification herein is taken  
in conjunction with the drawings appended hereto  
wherein:

15

Figures 1a and 1b show a general schematic  
diagram showing software and functional features



of a present invention item locator system,  
including the method of creating item /location  
data pairs;

Figure 2 shows a schematic diagram  
5 illustrating the physical functions of a present  
invention voice recognition item locator device  
after the item/location information data pairs  
have been created; and,

Figure 3 shows a schematic diagram of a  
10 present invention device illustrating details of  
a voice recognition submodule used therein.

#### DETAILED DESCRIPTION OF THE PRESENT INVENTION

The present invention is a voice  
activated/voice responsive item locator and  
15 system. By "item" is meant a place or thing that  
a user desires to locate. Thus, a item could be a

particular brand of canned string beans, a type  
of outdoor stain, a booth at a convention, a  
particular part in inventory for sale, assemblage  
or distribution, a particular automobile in a  
5 production facility lot or in a large parking  
garage, or a room, a functional group or a person  
in an office building or the like. The response  
may be in the form of a word or sentence  
presented visually or audibly and it may  
10 designate an aisle, a shelf, a bin number, a room  
number, a row and slot or space, etc.

The voice recognition system digitizes words  
spoken via a receiver (microphone) handset,  
headset, or built-in microphone for conversion  
15 from analog to digital utilizing a continuous  
speech recognition digital signal processor

(DSP). The main support structure may be a conventional type housing for phones and other communications devices, may be of a different shape or configuration or may be built into a device such as a wall or desk unit, with or without monitor. It may be portable or permanently affixed and could be powered by any means available, e.g. AC or DC current. In the portable mode, the system would be wireless for the user and would, in that respect operate like a cell phone, two way radio, "walkie talkie" or other short distance wireless device, but would have a processor at a central or fixed location having the same features as described above, i.e., the DSP with programming capabilities, etc.

The DSP is connected to a programmable

microprocessor and either by customized input or a standard program, the system enables the user to quickly enter voice-activated fields, e.g., such as "Where is...", "Find...", etc.

- 5 Verification of voice recognition accuracy (prior to execution) is optional and may be accomplished via synthesized voice playback and/or a screen confirmation which requires a "YES" or "NO" to execute or open for revision. In some preferred
- 10 embodiments, a screen, e.g., LCD, enables visual feedback during input phase, with support for deletion, insertion, correction, etc.
- Cancellation of the entire command or programming instructions may be possible at any time (prior
- 15 to execution), via keystroke or voice command.

Another important aspect of the present

invention is the inclusion into the system of software and hardware (equipment) to utilize a method of creating item location information for the system. It involves using item-identifying bar codes on items to be included and using location-identifying bar codes for corresponding locations which are also attached to the items (products). The location-identifying bar codes are typically physically placed on the products themselves at the store, warehouse, lot, etc. when convenient and after the location has been designated or determined. For example, when the products are located on aisles, shelves, bins, drawers, floor area grids, etc. and specific locations are known, appropriate personnel may create and physically attach the location-

identifying bar code label to the product.

The location-identifying bar codes may be custom created for the locations or may be established as a universal location system.

5 Alternatively, a manager could use existing UPC bar codes for the locations, provided that they were different from the items to be located, and provided that the system were programmed to correlate these particular codes to specified  
10 locations.

The item-identifying bar codes are typically located on the items themselves, but when more than one identical item is included, a single item of the set of identical items will be  
15 sufficient for the method to work. However, it is preferred that all items in each set have the bar

codes located thereon. In some preferred  
embodiments, the bar codes for the items are  
Universal Price Code (UPC) bar codes, but the  
present invention need not be limited thereto,  
5 such as when it would be more appropriate to  
create unique identifying codes for each and  
every item, such as automobiles, artwork, etc.

The essential features of the present  
invention involve the creation of a voice-based  
10 guide or locator and the creation of appropriate  
item/corresponding location data base, to offer  
enhanced convenience and speed to users for  
location of one or more items.

Figures 1a and 1b show a general schematic  
15 diagram of a present invention system showing  
general software features and functional

features. Thus, the present invention system includes a method, software and hardware for the creation of item/location data pairs, as described above. In Figure 1a, the basic aspects of the item/location information data creation are set forth in schematic form. The unique item-identifying bar codes are attached 2 to at least one of each different item for a plurality of sets of items, each set having items different from the items in the other sets. Likewise, unique corresponding location-identifying bar codes are attached 4 to the products, and, subsequently, they are read 6 in predetermined manner so that the program recognizes sequences and creates data pairs to develop the item/location vocabulary for the system. This



information is included in manager inputs 10  
(reference also Figure 1b). The method shown in  
Figure 1a is repeated as needed for updating 8.

Figure 1b illustrates other features of the  
5 present invention and includes a central  
processor 1 which may be an external or internal  
component, i.e., within a single unit or at a  
separate location from audio receivers and  
transmitters , e.g., microphones/speakers for  
10 user inputs and feedback to users.

The system may be preprogrammed with the  
user being required to follow concise  
instructions for activation and operation, or may  
be programmable to alter, add or enhance ease or  
15 methods of use, e.g. through a limited access  
code, for manager inputs 3 of user instructions.

In any event, manager inputs 3 shall include functional selections and inputs of items and their locations, with provision for subsequent access for modifications. This programming may include direct keyboard, voice, etc., and, as mentioned, may include security capabilities for preventing unauthorized use, e.g. voice identification (user recognition) or user security code system, as well as other options which may be included therein, such as a "help" detailed manager instruction section.

Once the system has been programmed for use, the user operation unit(s) provide functional access, which may be passive, i.e., the user speaks, picks up a phone, presses a button, or otherwise takes some action to

activate the system; or it may be active, i.e., a proximity sensor, a periodicity timer, or other internal mechanism may automatically activate the system and could trigger an audio or visual query, such as "May I help you locate a product?"

Once the system has been activated and a user has stated the necessary words of input to activate the device, recognition/non-recognition response 7 results from processing the user inputs to central processor 1 , and audio and/or video response unit(s) 9 provide feedback 11 to the user, either by answering the inquiry, conditionally defaulting, e.g., asking for a repeat or a restate the question, or fully defaulting, e.g. directing the user to a courtesy desk or check out counter for help.

Figure 2 shows a schematic diagram illustrating a present invention voice activated/voice responsive item locator system, showing the physical arrangement and function of components after the item/corresponding location information has been inputted. Thus, symbol 17 indicates an optional user prompter proximity sensor and symbol 21 is a microphone or equivalent component for voice input. The voice input is sent to audio controller 19 and to automatic speech recognition unit 23 and is converted from analog to digital signals. CPU/Memory 25 compares the digital signals to the set up or dictionary of digital words or phrases in memory. Once a match is found, the system processor 27 and data storage 31 operate to

respond with an answer or a default instruction  
or a query by providing digital text to text-to-  
speech generator 29, which provides audio  
feedback to a user via audio controller 19 and  
5 speaker 33. Feedback to a user may also be  
provided on visual screen 37 via display  
controller 35. Keyboard 39 is used for manager  
set up and modifications.

Figure 3 shows the details of one preferred  
10 embodiment of the submodule used in the present  
invention device. The voice recognition  
component converts an acoustic signal into a  
sequence of labels. The system takes the raw  
acoustic data, and processes it through the  
15 recognizer. The recognizer then matches it  
against a set of models using a decoder that

generates a recognition token. This token  
represents what the user said as either a single  
word or utterance. The recognizer itself does  
not interpret the meaning of the recognized  
5 output, that is the function of the interpreter  
(described later). The recognizer uses Hidden  
Markov Models (HMMs) to provide for a continuous  
speech recognition engine. HMMs do not process  
the acoustic signal directly but instead split  
10 the signal into a sequence of discrete  
observations. These observations are derived  
from a digital representation of the signal that  
had been converted from the analog signal  
generated by the microphone. During recognition,  
15 the likelihood of each model (or sequence of  
models) matching the incoming signal is

calculated. The recognizer simply selects the most likely model to decode the signal. As this is done continuously, the recognizer can process speech as opposed to isolated words, allowing the user to talk more naturally.

Each acoustic model represents a short sound. The interpreter combines these sounds into words using a dictionary. This dictionary specifies the pronunciation of each word in terms of the acoustic models. After identifying the most likely word, the interpreter then joins sets of models together (using a Viterbi decoder) in a series of pre-defined connections such that paths can be established to provide for a degree of "natural language" recognition; in other words, the user can say "Find hammers", "Where are

hammers" or "hammers" and they are all understood to mean the same thing. Moreover, these sets of models and dictionaries are interchangeable, allowing the same voice recognition component to be used in a variety of applications.

As the voice recognition component is running continuously, there needs to be a way to distinguish background conversations that might accidentally trigger an unwanted action by the device. For example, two people standing by a voice-activated device might be discussing locations of different goods in a supermarket and be misinterpreted or undesireably responded to. To avoid this problem, the recognition unit requires a command word to trigger before beginning further recognition. The trigger word



is a user-definable setting.

Thus, in Figure 3, initialization 51  
initiates monitoring 53 for a trigger word from a  
user. When a word is received, it is analyzed to  
5 determine whether or not a trigger word 55 has  
been received. If not, signal 57 returns the  
status to monitoring 53 for a new word. This  
loop continues until a trigger word is recognized  
and an inactivity timer 59 is started. The  
10 monitor 61 proceeds with the monitoring for the  
next word and waits for timer pop 65. When an  
event 63 is received, timer pop 65 returns to the  
monitor 53 to continue the monitoring process and  
the voice data is sent to interpretation 67. If  
15 it is understood 69, an action 75 is processed and  
feedback function 77 is performed. Additionally,

signal 79 prompts user 71. Likewise, if the interpretation is not understood 69, user 71 is prompted and via signal 73, timer 59 begins again. These cyclings operate on a continual basis while the system is initiated. Voice activation may also be used to shut down the system.

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.